Reliability of Strength Testing using the Advanced Resistive Exercise Device and Free Weights

Kirk L. English, M.A.[1]

James A. Loehr, M.S.[2]

Mitzi A. Laughlin, Ph.D.[2]

Stuart M.C. Lee, M.S.[2]

R. Donald Hagan, Ph.D.[3]

[1]JES Tech, Houston, TX

[2]Wyle Integrated Science and Engineering Group, Houston, TX

[3]NASA Johnson Space Center, Houston, TX

ABSTRACT

Introduction: The Advanced Resistive Exercise Device (ARED) was developed for use on the International Space Station as a countermeasure against muscle atrophy and decreased strength. This investigation examined the reliability of one-repetition maximum (1-RM) strength testing using ARED and traditional free weight (FW) exercise. *Methods*: Six males (180.8 ± 4.3 cm, 83.6 ± 6.4 kg, 36 ± 8 y, mean ± SD) who had not engaged in resistive exercise for at least six months volunteered to participate in this project. Subjects completed four 1-RM testing sessions each for FW and ARED (eight total sessions) using a balanced, randomized, cross-over design. All testing using one device was completed before progressing to the other. During each session, 1-RM was measured for the squat, heel raise, and deadlift exercises. Generalizability (G) and intraclass correlation coefficients (ICC) were calculated for each exercise on each device and were used to predict the number of sessions needed to obtain a reliable 1-RM measurement (G ≥ 0.90). Interclass reliability coefficients and Pearson's correlation coefficients (R) also were calculated for the highest 1-RM value (1-RM_{peak}) obtained for each exercise on each device to quantify 1-RM relationships between devices. *Results*:

	FW			ARED			Between Devices	
	G-coefficient from 3 sessions	G-coefficient from 4 sessions	# of sessions needed to obtain G ≥ 0.90	G-coefficient from 3 sessions	G-coefficient from 4 sessions	# of sessions needed to obtain G ≥ 0.90	R	ICC
Squat	0.91	0.94	3	0.86	0.89	5	0.56	0.71
Heel raise	0.89	0.94	3	0.90	0.94	3	0.33	0.50
Deadlift	0.99	0.99	1	0.97	0.98	1	0.17	0.29

Discussion: Three or fewer 1-RM sessions using ARED are required to obtain a reliable 1-RM measure for heel raise and deadlift exercises, but five sessions are needed to achieve reliable

values when performing 1-RM testing for the squat exercise. Reliable values were obtained using FW after three sessions for all three exercises. Neither FW nor ARED 1-RM accurately predicted the 1-RM measured using the other device.

ACRONYMS

0-g	Microgravity
1-g	Earth gravity
1-RM	One-repetition maximum
ANOVA	Analysis of Variance
ARED	Advanced Resistive Exercise Device
BMD	Bone Mineral Density
EVA	Extravehicular activity
FW	Free weights
iRED	interim Resistive Exercise Device
ISS	International Space Station
NASA	National Aeronautics and Space Administration
SE	Standard error of the mean

INTRODUCTION

Musculoskeletal deconditioning is apparent after sustained microgravity exposure. Significant muscle atrophy, change in muscle morphology, and reduction in strength have been found after spaceflights as short as 5 to 11 days (Edgerton, 1995; Greenisen et al., 1999). Following long-duration spaceflight, strength and muscle mass losses approach 30% in some lower body muscle groups despite the performance of inflight exercise countermeasures (Leblanc et al., 2000; Lee et al., 2004). Reduced muscle strength during long-duration spaceflight might increase fatigue and injury risk during strenuous tasks such as extravehicular activity (EVA) during International Space Station (ISS) construction and planetary exploration. Bone also undergoes deleterious changes such as increased calcium turnover during and reduced bone mineral density (BMD) following long-duration spaceflights (Smith et al., 1999). These changes might increase the risk of fracture in crew members during strenuous activities, falls, or upon re-exposure to gravitational stress (Lang, 2006).

Resistive exercise training is currently employed as a countermeasure to maintain muscle strength, muscle mass, and BMD during spaceflight. The interim Resistive Exercise Device (iRED), the current resistive exercise device on the ISS, is not as effective as ground-based FW training for increasing strength and BMD in ambulatory subjects (Schneider et al., 2003) and does not provide the high loads that have been shown to be protective during bed rest (Bamman et al., 1996; Shackelford et al., 2004). Results from early ISS missions suggest that iRED exercise may not be protective of isokinetic muscle strength (Lee et al., 2000) and BMD (Lang et al., 2004). Although iRED exercise was not the only countermeasure employed on these missions, perhaps the functional characteristics of iRED, such as a lower net eccentric force of

approximately 60-70% concentric force, a lack of inertial forces, and an altered resistance profile relative to FW training explain these differences (Lee et al., 2004).

In response to these limitations, the Advanced Resistive Exercise Device (ARED) was designed by the National Aeronautics and Space Administration (NASA) to provide forces that mimic lifting FW in normal gravity (1-g) and will replace iRED as the primary resistive exercise countermeasure device on the ISS. ARED incorporates three major improvements over iRED: greater maximal loading (275 kg peak load), improved eccentric:concentric ratio (eccentric load is approximately 90% of concentric), and simulation of inertial forces produced during free weight exercise.

Before replacing iRED as the primary resistive exercise device on the ISS, the NASA Exercise Physiology Laboratory will complete a training study comparing the physiologic adaptations during a 16-week resistive exercise training protocol using FW and ARED. The one-repetition maximum (1-RM), defined as the maximum amount of weight a person can lift one time, will be measured before, during, and after training. Healthy, untrained subjects are commonly recruited to participate in physiologic investigations to obtain a homogenous sample population responsive to resistive exercise training stimuli (Baechle 2002). A high-magnitude training response may be helpful when trying to characterize or evaluate the effects of a particular training program or increase the likelihood of detecting changes to training using particular exercise hardware. However, there are disadvantages to using untrained subjects in a strength training study, including their inherent unfamiliarity with proper lifting technique and discomfort while performing maximal efforts. To obtain a reliable measure of strength using the fewest number of

testing sessions and minimize potentially confounding factors, subjects must be familiarized with proper technique and strongly encouraged to exert themselves maximally. Pre-training strength tests that are too frequent or are greater in number than required may induce training effects that the investigator intends to measure and thus decrease the likelihood of detecting changes. Unpublished data from our laboratory suggest that, in untrained subjects, three 1-RM sessions are required to obtain a reproducible 1-RM using FW (James Loehr, personal communication), but no comparable data are available for ARED.

The primary purpose of this study was to quantify the number of testing sessions required to obtain a reliable 1-RM measurement using ARED in untrained subjects who have similar characteristics to those who will participate in the subsequent training study. The secondary objective of this project was to determine whether there are systematic differences between 1-RM strength measures using FW and ARED.

METHODS

Subjects

Six males (180.8 ± 4.3 cm, 83.6 ± 6.4 kg, 36 ± 8 y, mean ± SD) volunteered to participate in this project. Subjects were required to have abstained from any resistive exercise training in the previous six months and to pass a modified Air Force Class III physical examination. Subjects received written and verbal explanations of the testing protocols and provided written informed consent. The test protocols and procedures were reviewed and approved by the NASA Johnson Space Center Committee for the Protection of Human Subjects.

Subjects performed four 1-RM testing sessions each for FW and ARED (eight total testing sessions) using a balanced, randomized, cross-over design. During each session, 1-RM was measured for three exercises: squat, heel raise, and deadlift; the order of exercises performed was held constant across sessions. Subjects completed all four sessions on one device before proceeding to testing on the other. Each session was separated by 5-10 days.

Equipment

FW squat was conducted using a standard Olympic barbell and plates and a York Barbell half rack (York, PA). FW heel raise was performed on a Smith machine (Bigger Faster Stronger Model 300052, Salt Lake City, UT). FW deadlift was performed using a standard Olympic barbell and plates. All ARED squat, heel raise, and deadlift were performed using ARED.

The primary resistive mechanism in ARED (Figure 1) is a pair of vacuum cylinders, each containing a large piston. Resistance is provided by the movement of the pistons within the vacuum of the cylinders. The piston rods are attached to the arm base assembly which acts as a lever arm when the bar is moved. Load is increased or decreased by the load adjustment handle which pivots the arm base assembly, either lengthening or shortening the resistance arm. In addition, flywheels mounted below the vacuum cylinders are rotated as the bar and arm base assembly move, providing an inertial resistance that the user must overcome when accelerating or decelerating the bar during exercise. These were added to mimic the inertial forces experienced during exercise using FW.

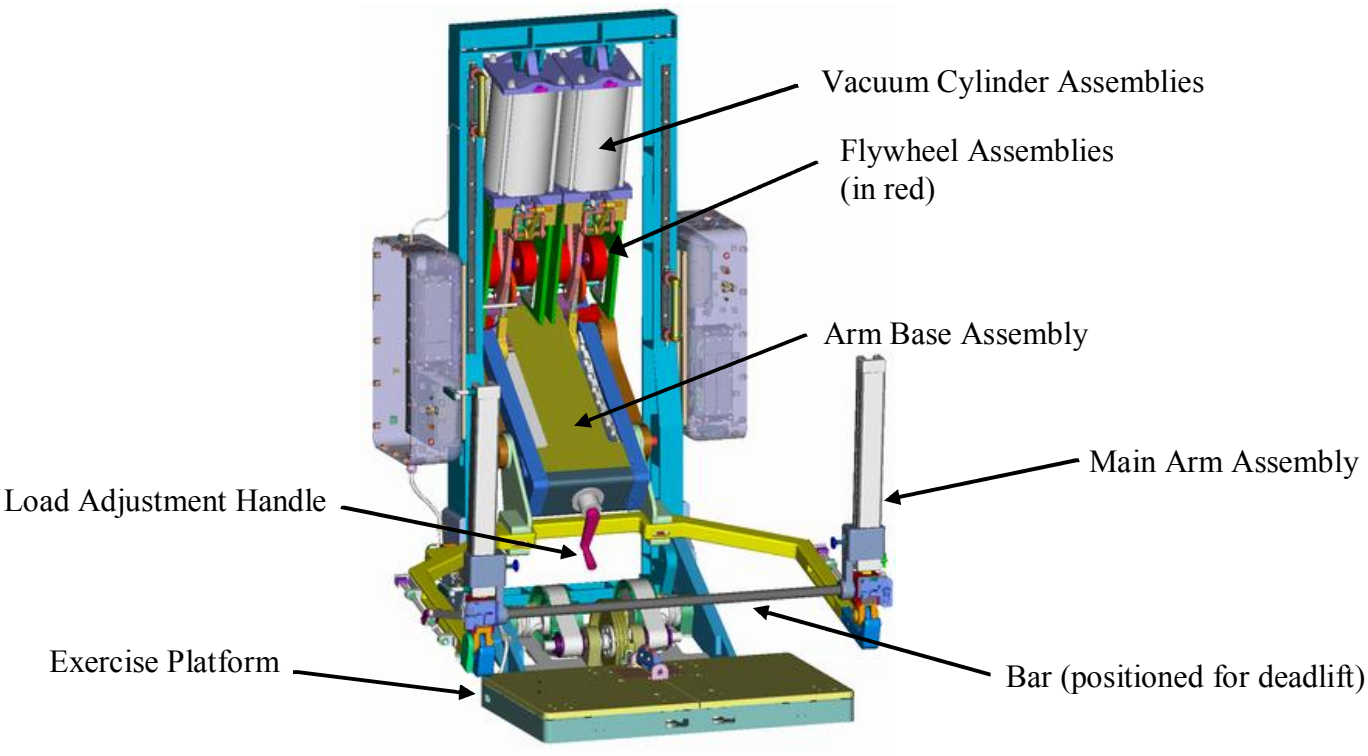

Figure 1. Schematic of ARED

Procedures

Subjects were instructed in proper lifting technique and were spotted at all times by three qualified trainers. The testing protocol included warm-up sets for each exercise and progressed in a defined manner to incrementally higher loads until the subject could not successfully complete the lift in the defined range of motion or utilized poor lifting technique. Subjects were permitted two successive failures at a given load before testing was terminated. The 1-RM value was determined as the highest load successfully lifted. No previous 1-RM values existed for these subjects prior to the first testing session. Therefore, during the first session, conservative loads and increments for successive attempts were utilized in a progression similar to that used in subsequent testing sessions (Table 1). Load and repetition progressions for sessions 2-4 were

based upon a percentage of the 1-RM measured in the previous testing session (Table 1). Load increments above a subject's previous 1-RM were selected by the test operators based on their perception of the difficulty the subject experienced in completing the previous set. For all exercises, subjects were instructed to complete a single repetition by lowering the weight for approximately two seconds and raising it as quickly as they wished.

Table 1. 1-RM Testing Protocol

Set #	Reps	% of previous 1-RM (load)
1	8	50 %
2	8	60 %
3	5	70 %
4	3	80 %
5	1	90 %
6	1	100 %
7	1	100+ %
8	1	100+ %

Squat Exercise. During the first session with each device, squat depth was determined as the vertical distance the bar traveled from the standing position to a stance at which the femur was parallel to the floor. The target depth for the squat was held constant across all subsequent testing sessions within and across exercise devices and was measured by a linear encoder (Ergotest Technology, Langesund, Norway) that was attached to the bar and interfaced with a laptop computer (Compaq Armada E500, Houston, TX) using a customized LabView software program (National Instruments, Austin, TX). When subjects reached the required depth, they

received a computer-generated audio cue to reverse direction of movement and attempt to raise the load to the starting, or standing, position (Figure 2).

Figure 2. Squat start/finish and mid-position using ARED

Heel Raise Exercise. The forefeet were positioned on a wooden toeboard (8.5 cm high x 9 cm wide x 59 cm long) to allow for ankle dorsiflexion. To begin each set, two spotters assisted the subjects in raising the bar so that they began each set with maximum ankle plantarflexion (Figure 3). Subjects then dorsiflexed their ankles while keeping their knees and hips straight until their heels touched the ground. Upon ground contact, subjects immediately extended their ankles to maximum plantarflexion. During the first session, the bar height at peak plantarflexion for each subject was determined to the nearest full centimeter using a linear encoder attached to the bar. A successful lift was confirmed with the linear encoder as the computer generated an audio cue when the subject attained peak plantarflexion. Peak plantarflexion bar height was held constant between exercise devices.

Figure 3. Heel raise start/finish and mid-position using ARED

Deadlift Exercise. A successful deadlift was defined as raising the bar from its lower most, resting position to the subject's upper thigh, with the hips and knees fully extended and the scapulae retracted (Figure 4). Subjects were also required to maintain a flat or arched back throughout the lift. Subjects used an alternating grip with the hands positioned outside the legs. Test operators judged whether or not a subject had successfully completed a given repetition using correct technique. ARED deadlift was conducted with the starting point as the lowest possible bar setting. To replicate the ARED bar height and ensure an identical starting position on both devices, two dense foam pads were placed beneath the FW plates, elevating the bar 2 cm.

Figure 4. Deadlift start/finish and mid-position using ARED

Statistical Analyses

One-way repeated measures ANOVA and Holm-Sidak *post-hoc* tests were used to detect differences between sessions within devices. Data are expressed as mean ± SE. Statistical significance was defined *a priori* as $p \leq 0.05$. Generalizability (G) and intraclass correlation coefficients (ICC) were calculated for each exercise by device and were used to determine the number of sessions needed to obtain a reliable 1-RM measurement; G-coefficients ≥ 0.90 were considered representative of reliable 1-RM values (Shavelson, 1991). If a G-coefficient ≥ 0.90 was not attained in four sessions, the number of sessions needed to reach that level was calculated. Interclass reliability coefficients and Pearson's correlation coefficients (R) also were calculated for the highest 1-RM value (1-RM_{peak}) obtained for each exercise on each device to quantify 1-RM relationships between devices. Generalizability, reliability, and correlation were considered excellent if ≥ 0.95, high if ≥ 0.90, good if ≥ 0.80, fair if ≥ 0.70, poor if ≤ 0.70, and very poor if ≤ 0.40.

RESULTS

There were no significant differences in 1-RM between sessions within each device for squat (Figure 5). Similarly, no significant differences between sessions were seen for FW heel raise, but ARED heel raise 1-RM values for Sessions 2-4 were significantly greater than the first session (Figure 5). All FW deadlift sessions were significantly different from each other. For ARED deadlift, 1-RM values for Sessions 2-4 were significantly greater than Session 1 and Session 4 was significantly greater than Session 2 (Figure 5).

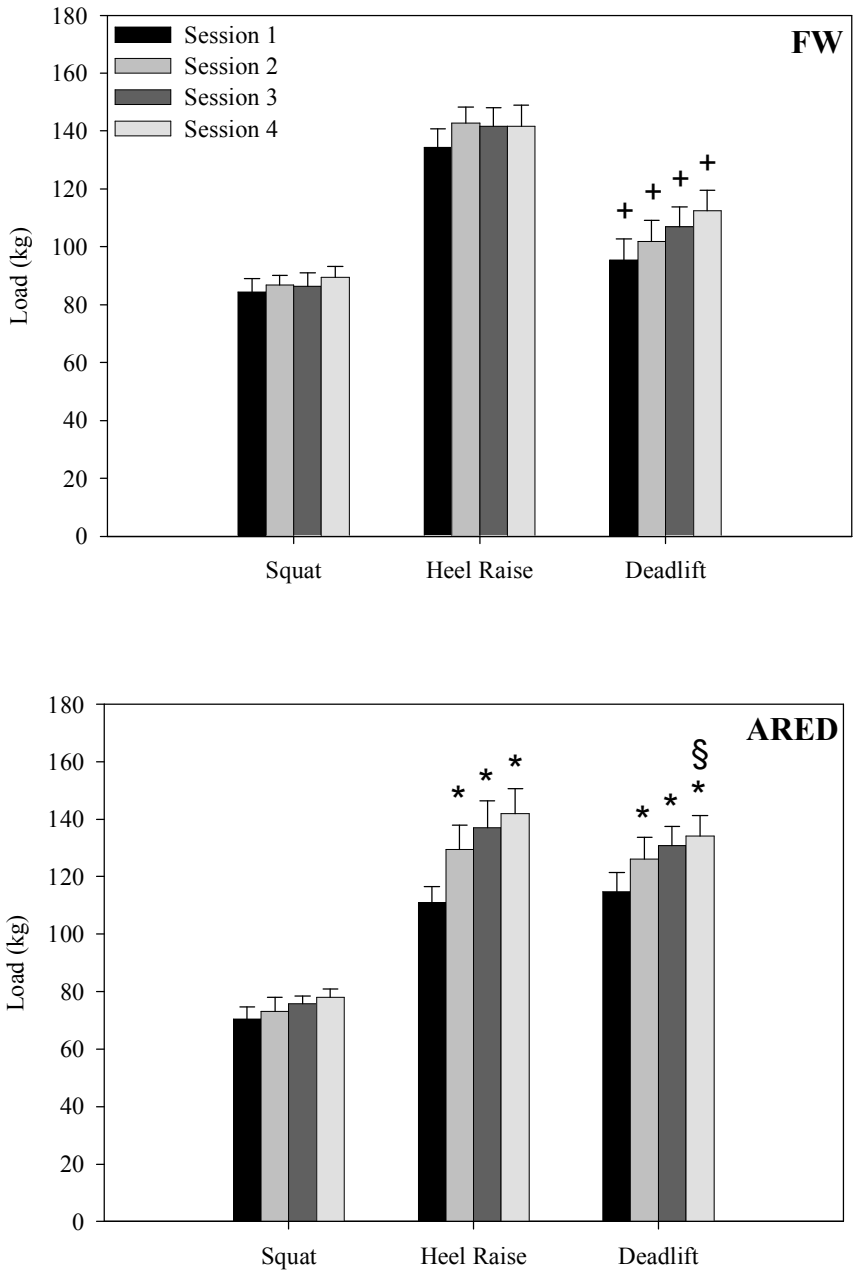

Figure 5. Within device differences in mean (± SE) 1-RM between sessions for squat, heel raise, and deadlift. + all within device pairwise comparisons significantly different ($p < 0.05$), * significantly different from Session 1 ($p < 0.05$), § significantly different from Session 2 ($p < 0.05$).

Generalizability coefficients for all four sessions were high for FW squat, good for ARED squat, and high for both FW and ARED heel raise (Table 2). G-coefficients for deadlift were in the excellent range for both devices. Similarly, intraclass correlation for FW squat was high and ARED squat was considered good. Intraclass correlation coefficients were excellent for heel raise and deadlift for both devices.

Table 2. G-coefficients and intraclass correlation coefficients for FW and ARED.

	FW				ARED			
	G-coefficient for 3 sessions	G-coefficient for 4 sessions	# of sessions needed to attain G ≥ 0.90	ICC (3,1)	G-coefficient for 3 sessions	G-coefficient for 4 sessions	# of sessions needed to attain G ≥ 0.90	ICC (3,1)
Squat	0.91	0.94	3	0.92	0.86	0.89	5	0.83
Heel Raise	0.89	0.94	3	0.95	0.90	0.94	3	0.99
Deadlift	0.99	0.99	1	0.99	0.97	0.98	1	0.99

Pearson's correlation coefficients for 1-RM_{peak} values between devices was poor for squat and extremely poor for heel raise and deadlift; interclass reliability was fair for squat, poor for heel raise, and extremely poor for deadlift (Table 3).

Table 3. Pearson's correlation coefficients and interclass reliability of 1-RM_{peak} for squat, heel raise, and deadlift on FW and ARED.

	R	ICC (FW vs. ARED)
Squat	0.56	0.71
Heel Raise	0.33	0.50
Deadlift	0.17	0.29

DISCUSSION

The purpose of this investigation was to determine: 1) how many sessions are needed to obtain a reproducible 1-RM measurement in untrained subjects using FW and ARED and 2) whether 1-RM measurements obtained on either FW or ARED can be used as 1-RM values on the other device.

We found that three 1-RM sessions using ARED provide reliable 1-RM values for heel raise and deadlift. However, five 1-RM sessions are needed for squat using ARED. FW required three, three, and one session(s) to obtain reliable 1-RM measures for squat, heel raise, and deadlift, respectively.

In untrained subjects, there was only fair to extremely poor reliability of 1-RM values obtained using the two devices for the exercises tested in this study. It is unlikely that either FW or ARED 1-RM values can be used to accurately predict 1-RM measures on the other device.

Reproducibility of 1-RM measures

Previous studies have shown that untrained subjects require multiple sessions to obtain reproducible 1-RM values, and that age, sex, exercise, and type of device also may influence this. For example, Ploutz-Snyder and Giamis (2001) determined that untrained young adult females needed 2-5 sessions (mean of 3.6) to obtain a reproducible 1-RM when performing seated bilateral knee extensions. Similarly, Phillips et al. (2004) reported reliable 1-RM measurements for machine bench press and leg press using untrained seniors following three familiarization sessions and 2-3 actual 1-RM sessions. In contrast, Schoeder et al. (2007) found no more than a 2.3 % change between the first two 1-RM sessions using untrained elderly males performing chest press, latissimus pull-down, leg press, and leg flexion. Although utilizing untrained subjects, these studies all employed exercise machines which might require less familiarization than FW or ARED due to their restricted movement planes which require less neuromuscular coordination (McCaw, 1994).

Besides the subject population, a major difference between the aforementioned studies and the present investigation is the exercise device utilized. ARED is an exercise machine, but does not restrict bar motion to a single trajectory. However, unlike FW exercise in which six degrees of freedom (vertical, fore-aft, and horizontal translation, and roll, pitch, and yaw rotation) are possible for the bar, the ARED bar only allows vertical and horizontal translation as a rotation about a pivot point. Because bar motion is more fixed on ARED than FW, there may be decreased neuromuscular coordination required to perform a successful lift using ARED. Because control of the body and external resistance is a critical task when performing a lift, lifting on ARED may therefore be different than lifting on FW, resulting in a decreased number

of sessions required to obtain a reproducible 1-RM value. Examination of the G-coefficients for heel raise and deadlift suggests that for these two exercises, the learning curve is actually similar for ARED and FW; however, for the squat exercise, subjects needed more sessions to obtain a reproducible 1-RM value when using ARED than when using FW.

Using ARED, the squat exercise required five sessions to obtain a reproducible value as opposed to the other two exercises which required three or fewer sessions. This is perhaps due to differences in the complexity of the squat mechanics which made it a more difficult movement to learn. During the descent, subjects reported a loss of balance on ARED, because of a rearward force due to the limitation in bar range of motion. However, adjusting foot position, increasing trunk flexion, and performing multiple sessions reduced this perception in individual subjects. When squatting on FW, some subjects reported a similar sensation which was also resolved using the above-mentioned techniques. However, it was clear that subjects required more sessions to successfully implement these corrective steps when using ARED than FW. Also, when squatting on ARED, at the bottom of the lift subjects often stated that they felt like they were being "folded over." Although spotters took special care to coach subjects to prevent this, this sensation was most likely due to an apparent increase in trunk flexion. It is unclear, however, whether the peculiarities associated with ARED squat technique will persist in microgravity.

While the results of this study may guide the number of sessions needed for untrained subjects in future studies, it is possible that fewer sessions will be required to obtain a reliable measure of strength in trained subjects, such as astronauts preparing for spaceflight who have previous

experience with the testing hardware. For example, Doan et al. (2002) reported no significant differences between three 1-RM bench press measures in trained male subjects. In another study utilizing trained males, Rhea et al. reported a high interclass correlation (R = 0.99) between the second and third of three 1-RM sessions for bench press and leg press (2002). Subjects did however, complete six "instruction/training sessions" prior to the three 1-RM tests (Rhea et al, 2002). In the present study, instruction was incorporated into the actual 1-RM sessions which makes comparison with the data of Rhea et al. difficult. We suspect that astronauts preparing for spaceflight who have previous experience with ARED will not require as many sessions as an untrained subject, but this has yet to be tested.

Relationship between FW and ARED 1-RM

Although the relationship between 1-RM measured using the two devices for the squat was moderate across sessions, eliciting a correlation coefficient of 0.71 (fair), this was not true for the other exercises despite the fact that they required fewer sessions to obtain reproducible values. The interclass correlation coefficient for heel raise was much lower (ICC (3,1) = 0.50). This poor reliability is likely due to inherent differences in the mechanics provided by FW and ARED when performing these exercises. FW heel raise was conducted on a Smith machine which has only one degree of freedom and provides more stability to the test subject compared to using a traditional barbell. Greater stability not only provides increased subject safety, but also allows greater maximal force production (Anderson, 2004). Training loads are based on 1-RM values, thus, their determination under conditions of increased stability ultimately provides a greater training stimulus due to the higher 1-RM values thus elicited. During heel raise on ARED, movement is restricted in the frontal plane, but subjects still had to balance themselves to prevent

from falling forward or backward (two degrees of freedom). This contrasts with the stability provided in both the frontal and sagittal planes when performing heel raises on a Smith machine. Thus, ARED 1-RM values may have been affected by the lifters' reduced stability.

Deadlift had the lowest interclass reliability coefficient (0.29). One possible explanation for this is an inherent difference in the mechanics of FW and ARED deadlift. During FW deadlift, the bar travels in a vertical direction (Escamilla, 2000) due to the gravitational force on the mass. In contrast, ARED has a pivot point where the bar attaches to the main arm assembly, providing a variety of possible resultant force vectors. Due to ARED's construction, subjects were able to lean backwards during the deadlift without concern for losing their balance. The resultant friction of the subjects' shoes created a braking force which provided added stability and enabled subjects to generate greater forces. Although operators watched for this and instructed subjects not to lean back, the differences in 1-RM values between FW and ARED suggest that some subjects still employed this or some alternate technique. Future research to characterize biomenchanical and kinematic differences between FW and ARED lifting is needed.

Conclusions

The intraclass correlation for a given exercise on a particular device (either FW or ARED) ranged from good to excellent. G-coefficients of ≥ 0.90, our indice of reproducible 1-RM values, were obtained using FW after three sessions for squat and heel raise and one session for deadlift. In comparison, subjects required five, three, and one session(s) to attain a G-coefficient ≥ 0.90 for the same three exercises on ARED, respectively. Thus, ARED squat was the only exercise that fell outside of our previously established laboratory standard of three pre-training 1-

RM sessions for untrained subjects. To effectively prescribe exercise or establish baseline values for untrained subjects, three 1-RM sessions using ARED are needed to provide reliable 1-RM values for heel raise and deadlift. However, five 1-RM sessions are needed to obtain reliable 1-RM values for squat using ARED. Crew members and other trained subjects may need fewer sessions to obtain reliable 1-RM values because of their previous familiarity with the exercises, but this remains to be tested.

We also sought to determine if a relationship existed between 1-RM values measured using FW and using ARED. In untrained subjects, there was only fair to extremely poor reliability of 1-RM values obtained using the two devices for the exercises tested in this study. Crew members will perform resistive exercise using ARED on the ISS; their inflight strength training programs will be based on 1-RM values. These values either must be measured directly by completing a 1-RM test using ARED or predicted from FW 1-RM values. Based on the poor correlation coefficients, it is unlikely that either FW or ARED 1-RM values can be used to accurately predict 1-RM measures on the other device. Thus, crew members will be best served with considerable familiarization, training, and testing time using ARED prior to flight.

Acknowledgements:

We would like to thank our subjects for their enthusiastic participation in this study, Jason Bentley and Mark Leach for their extensive involvement with data collection and hardware configuration, and the ARED Engineering Team for their collaboration and cooperation. This report is dedicated to our friend and mentor, Don Hagan.

REFERENCES

1. Anderson KG and DG Behm. Maintenance of EMG activity and loss of force output with instability. *J Strength Cond Res.* 18(3): 637-640, 2004.
2. Baechle TR and RW Earle. Essentials of Strength Training and Conditioning. Champaign, IL: Human Kinetics, 2002.
3. Bamman MM, MS Clarke, DL Feeback, RJ Talmadge, BR Stevens, SA Lieberman, and MC Greenisen. Impact of resistance exercise during bed rest on skeletal muscle sarcopenia and myosin isoform distribution. *J Appl Physiol.* 84: 157-63, 1998.
4. Doan BK, RU Newton, JL Marsit, NT Triplett-McBride, LP Koziris, AC Fry, and WJ Kraemer. Effects of increased eccentric loading on bench press 1RM. *J Strength Cond Res.* 16(1): 9-13, 2002.
5. Edgerton VR, MY Zhou, Y Ohira, H Klitgaard, B Jiang, G Bell, B Harris, B Saltin, PD Gollnick, RR Roy, MK Day, and M Greenisen. Human fiber size and enzymatic properties after 5 and 11 days of spaceflight. *J Appl Physiol.* 78(5): 1733-39, 1995.
6. Escamilla RF, AC Francisco, GS Fleisig, SW Barrentine, CM Welch, AV Kayes, KP Speer, and JR Andrews. A three-dimensional biomechanical analysis of sumo and conventional style deadlifts. *Med Sci Sports Exerc*. 32(7): 1265-1275, 2000.
7. Greenisen MC, JC Hayes, SF Siconolfi, and AD Moore. Functional performance evaluation. Extended duration orbiter medical project final report, 1989-1995. Houston, TX: NASA Johnson Space Center; 1999: Section 3, 1-24.
8. Lang T, A LeBlanc, H Evans, Y Lu, H Genant, and A Yu. Cortical and trabecular bone mineral loss from the spine and hip in long-duration spaceflight. *J Bone Miner Res.* 19: 1006-12, 2004.
9. Lang T. What do we know about fracture risk in long-duration spaceflight? *J Musculoskelet Neuronal Interact*. 6(4): 319-321, 2006.
10. Leblanc A, C Lin, L Shackelford, V Sinitsyn, H Evans, O Belichenko, B Schenkman, I Kozlovskaya, V Oganov, A Bakulin, T Hendrick, and D Feedback. Muscle Volume, MRI relaxation times (T2) and body composition after space flight. *J Appl Physiol.* 89: 2158-2164, 2000.
11. Lee SMC, ME Guilliams, SF Siconolfi, MC Greenisen, SM Schneider, LC Shackelford. Concentric strength and endurance after long duration spaceflight. *Med Sci Sports Exerc*. 32: S363, 2000.
12. Lee SMC, K Cobb, JA Loehr, D Nguyen, and SM Schneider. Foot-ground reaction force during resistive exercise in parabolic flight. *Aviat Space Environ Med.* 75: 405-12, 2004.
13. McCaw ST and JJ Friday. A comparison of muscle activity between a free weight and machine bench press. *J Strength Cond Res*. 8(4): 259-264, 1994.
14. Phillips WT, AM Batterham, JE Valenzuela, and LN Burkett. Reliability of maximal strength testing in older adults. *Arch Phys Med Rehabil*. 85: 329-34, 2004.

15. Ploutz-Snyder LL and EL Giamis. Orientation and familiarization to 1RM strength testing in old and young women. *J Strength Cond Res.* 15(4): 519-23, 2001.
16. Rhea MR, SD Ball, WT Phillips, and LN Burkett. A comparison of linear and daily undulating periodized programs with equated volume and intensity for strength. *J Strength Cond Res.* 16(2):250-55, 2002.
17. Schneider SM, WE Amonette, K Blazine, J Bentley, SMC Lee, JA Loehr, AD Moore Jr., M Rapley, ER Mulder, and SM Smith. Training with the International Space Station Interim Resistive Exercise Device. *Med Sci Sports Exerc.* 35: 1935-45, 2003.
18. Schroeder ET, Y Wang, C Castaneda-Sceppa, G Cloutier, AF Vallejo, M Kawakubo, NE Jensky, S Coomber, SP Azen, and FR Sattler. Reliability of maximal voluntary muscle strength and power testing in older men. *J Gerontol*. 62A(5): 543-49, 2007.
19. Shackelford LC, AD LeBlanc, TB Driscoll, HJ Evans, NJ Rianon, SM Smith, E Spector, DL Feeback, and D Lai. Resistance exercise as a countermeasure to disuse-induced bone loss. *J Appl Physiol.* 97: 119-29, 2004.
20. Shavelson RJ and NM Webb. *Generalizability Theory: a primer*. Sage Publications Newbury Park, CA, 1991.
21. Smith SM, ME Wastney, BV Morukov, IM Larina, LE Nyquist, SA Abrams, EN Taran, C Shih, JL Nillen, JE Davis-Street, BL Rice, and HW Lane. Calcium metabolism before, during, and after a 3-mo spaceflight: kinetic and biochemical changes. *Am J Physiol.* 277: R1-10, 1999.